Choice Models in Operations

Lecture 3 : ARSP Remarks & Introduction to RUM Models

Instructor: Srikanth Jagabathula *Scribe: Ryan McNellis*

1 ARSP Remarks

1.1 Extension of Choice Rules

Define $\mathcal{S}_{\text{pairs}}$ to be the pairwise collection of subsets of $\mathcal{N}$:

$$\mathcal{S}_{\text{pairs}} = \{\{i,j\} : 1 \le i,j \le n, i \ne j\}$$

Suppose we specify the choice probabilities for the collection $\mathcal{S}_{\text{pairs}}$. Let p_{ij} denote the probability that i is preferred over j, i.e. $p_{ij} = \mathbb{P}(i|\{i,j\})$. Note that $p_{ij} + p_{ji} = 1$.

Definition 1 (Extension of a choice rule) *Given a stochastic choice function on the collection $\mathcal{S}$, an extension to a "bigger" collection $\mathcal{S}'$ is a choice function $\mathbb{P}'(\cdot|\cdot)$ such that $\mathbb{P}'(x|\mathcal{B}) = \mathbb{P}(x|\mathcal{B})\ \forall \mathcal{B} \in \mathcal{S}$.*

Theorem 1 *Suppose ARSP is satisfied for the collection $\mathcal{S}_{pairs}$. Then, there exists an extension of the choice rule such that it is stochastically rational for all possible subsets.*

Consider extending the choice function on $\mathcal{S}_{\text{pairs}}$ when $n = 3$. Note that one cannot simply define π by multiplying the pairwise choice probabilities and renormalizing, e.g.:

$$\pi(1 \succ 2 \succ 3) = \frac{p_{12}p_{13}}{z}, \text{where z is a normalizing constant.} \tag{1}$$

If this extension were valid, then $\pi(1 \succ 2 \succ 3) + \pi(1 \succ 3 \succ 2) + \pi(3 \succ 1 \succ 2) = p_{12}$. However, if we define π according to (1), then this does not necessarily hold:

$$\begin{aligned}
\pi(1 \succ 2 \succ 3) + \pi(1 \succ 3 \succ 2) + \pi(3 \succ 1 \succ 2) &= \frac{p_{12}p_{23}}{z} + \frac{p_{13}p_{32}}{z} + \frac{p_{31}p_{12}}{z} \\
&= \frac{p_{12}(p_{23} + p_{31}) + p_{13}p_{32}}{z} \\
&\ne p_{12} \text{ in general}
\end{aligned}$$

However, given $p_{ij} = \mathbb{P}(i|\{i,j\})$ over $\mathcal{S}_{\text{pairs}}$, the extension $\pi(\succ)$ can be approximated for any $\succ \in \mathcal{S}_n$ through Monte Carlo simulation. Simply compute the fraction of times the ordering $\succ$ appears when using the pairwise probabilities p_{ij} to "generate" valid permutations in $\mathcal{S}_n$. A valid permutation can be generated as follows:

1. $\forall\, i,j : 1 \le i < j \le n$, use p_{ij} to generate the pairwise relationship between i and j, i.e. sample $i \succ j$ with probability p_{ij}, else sample $j \succ i$.

2. Check that the generated pairwise relations satisfy transitivity. If so, then we have generated a valid permutation. Else, return to 1 and try again.

1.2 Geometric Interpretation of ARSP

We now interpret ARSP geometrically, focusing specifically on the collection $\mathcal{S}_{\text{pairs}}$.

Suppose the stochastic choice funciton $(p_{ij})_{1\le i,j\le n, i\ne j}$ is stochastically rational. Then, there exists a probabiliy mass function $\pi(\cdot)$ over preference lists such that

$$p_{ij} = \sum_{\succ \in \mathcal{S}_n : i \succ j} \pi(\succ) = \sum_{\succ \in \mathcal{S}_n} \pi(\succ)\mathbb{I}[i \succ j] \,.$$

We can write this more compactly using the same technique as in last lecture. Let $\underline{p}$ denote the vector representation of p_{ij}. For $n = 2$,

$$\underline{p} = \begin{bmatrix} p_{12} \\ p_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \pi(1 \succ 2) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \pi(2 \succ 1) \,.$$

Since $\pi(1 \succ 2)$ and $\pi(2 \succ 1)$ are non-negative and add up to one, $\underline{p}$ is the convex combination of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. In other words, $\underline{p}$ belongs to the convex hull of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

More generally, define $\underline{p_{\succ}}$ as a vector which encodes all possible pairwise relations with respect to $\succ$, i.e. $p_{\succ_{ij}} = \mathbb{I}[i \succ j] \quad \forall 1 \le i, j \le n, i \ne j$. Then, the pairwise probability vector $\underline{p}$, when rational, belongs to the convex hull of the vectors $\{\underline{p_{\succ}} : \succ \in \mathcal{S}_n\}$:

$$\underline{p} = \sum_{\succ \in \mathcal{S}_n} \underline{p_{\succ}} \pi(\succ)$$

Define the polytope $\mathcal{D}$ as follows:

$$\mathcal{D} = \left\{ \underline{x} \in \mathbb{R}^{n(n-1)\times 1} : \underline{x} = \sum_{\succ \in \mathcal{S}_n} \underline{p_{\succ}} \pi_{\succ} ; \sum_{\succ \in \mathcal{S}_n} \pi_{\succ} = 1, \pi_{\succ} \ge 0 \;\; \forall \succ \in \mathcal{S}_n \right\}$$

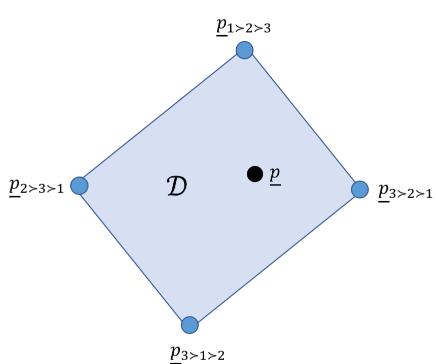

Figure 1: $\mathcal{D}$ corresponding to $n = 3$. The vertices of the polytope belong to the set $\{\underline{p_{\succ}} : \succ \in \mathcal{S}_n\}$ (note: only four of the vertices are included in the figure). Since $\underline{p} \in \mathcal{D}$, $\underline{p}$ is stochastically rational.

$\mathcal{D}$ thus represents the convex hull of $\{\underline{p}_{\succ} : \succ \in \mathcal{S}_n\}$. We know that if $\underline{p}$ is stochastically rational, then $\underline{p} \in \mathcal{D}$. Clearly, the converse is true as well: if $\underline{p}$ not stochastically rational, then there does not exist a feasible p.m.f. π over the preference lists, i.e. there does not exist a convex combination of the vectors $\{\underline{p}_{\succ} : \succ \in \mathcal{S}_n\}$ which equals $\underline{p}$. Thus,

$$\underline{p} \text{ is stochastically rational} \Leftrightarrow \underline{p} \in \mathcal{D}\,.$$

We can use this interpretation to easily show the equivalency between the ARSP condition and stochastic rationality:

$$\underline{p} \text{ is stochastically rational} \Leftrightarrow \underline{p} \text{ satisfies ARSP}$$

"$\Rightarrow$": Suppose that $\underline{p}$ is stochastically rational. It follows that $\underline{p} \in \mathcal{D}$. For any $\underline{z} \in \mathbb{R}^{n(n-1)\times 1}$, consider the following optimization problem:

$$\max_{\underline{x}} \underline{z}^T \underline{x} \qquad s.t. \quad \underline{x} \in \mathcal{D}$$

Because this is a linear programming problem, the optimal solution occurs at an extreme point, i.e. $\underline{p}_{\succ^*}$ for some preference list $\succ^* \in \mathcal{S}_n$. Also, because $\underline{p} \in \mathcal{D}$, $\underline{p}$ is a feasible solution to the LP. Thus, for all $\underline{z} \in \mathbb{R}^{n(n-1)\times 1}$,

$$\underline{z}^T \underline{p} \leq \underline{z}^T \underline{p}_{\succ^*} = \max_{\succ \in \mathcal{S}_n} \underline{z}^T \underline{p}_{\succ}\,.$$

This is simply the ARSP condition applied to the collection $\mathcal{S}_{\text{pairs}}$.

"$\Leftarrow$": Suppose $\underline{p}$ is not rational. Then, $\underline{p} \notin \mathcal{D}$. It follows that there exists a separating hyperplane between the point $\underline{p}$ and polytope $\mathcal{D}$, i.e. there exists a $\underline{z} \in \mathbb{R}^{n(n-1)\times 1}$ such that $\underline{z}^T \underline{p} > 0$ and $\underline{z}^T \underline{x} \leq 0\ \forall \underline{x} \in \mathcal{D}$. In particular, the latter condition implies that $\underline{z}^T \underline{p}_{\succ} \leq 0\ \forall \succ \in \mathcal{S}_n$.

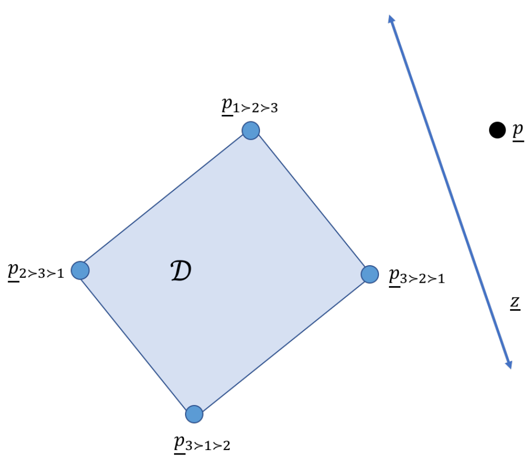

Figure 2: $\mathcal{D}$ corresponding to $n = 3$. $\underline{p}$ is not stochastically rational, and thus $\underline{p} \notin \mathcal{D}$. This implies the existance of a separating hyperplane $\underline{z}$.

Combining these inequalities, we arrive at

$$\underline{z}^T \underline{p} > \max_{\succ \in \mathcal{S}_n} \underline{z}^T \underline{p}_{\succ} ,$$

i.e., $\underline{p}$ does not satisfy ARSP.

2 Random Utility Maximization (RUM) Models

Suppose $\mathcal{N}$ is finite. Then, any probability mass function over preference lists can be represented by a distribution over utility vectors.

Suppose we are given $\pi : \mathcal{S}_n \to [0, 1]$. Let $\underline{\mathsf{U}} \in \mathbb{R}^n$ be a random utility vector corresponding to the n products. Then, we can easily construct a distribution $F_{\underline{\mathsf{U}}}(\cdot)$ which is consistent with π.

First, for all preference lists $\succ \in \mathcal{S}_n$, find a utility function $u_{\succ} : \mathcal{N} \to \mathbb{R}$ which represents $\succ$, i.e. $u_{\succ}(i) > u_{\succ}(j) \Leftrightarrow i \succ j$. Let $\underline{u}_{\succ}$ be the vector representation of $u_{\succ}(\cdot)$ such that the i^{th} component of $\underline{u}_{\succ}$ is $u_{\succ}(i)$. Then, define $F_{\underline{\mathsf{U}}}(\cdot)$ as follows:

$$\begin{cases} \mathbb{P}_F(\underline{\mathsf{U}} = \underline{u}_{\succ}) = \pi(\succ) & \text{for any } \succ \in \mathcal{S}_n, \text{ and} \\ \mathbb{P}_F(\underline{\mathsf{U}} = \underline{y}) = 0 & \text{when } \ \underline{y} \neq \underline{u}_{\succ} \ \text{ for any } \succ \in \mathcal{S}_n \end{cases}$$

Clearly, this distribution over utility vectors is consistent with the probability mass function π.

On the other hand, suppose we are given a distribution $F_{\underline{\mathsf{U}}}(\cdot)$ over utility vectors. This distribution induces a probability mass function over rankings, given by

$$\mathbb{P}(1 \succ 2 \succ \cdots \succ n) = \mathbb{P}(\mathsf{U}_1 > \mathsf{U}_2 > \cdots > \mathsf{U}_n)$$

For example, in the case of two products, we might be given a distribution $F_{\underline{\mathsf{U}}}(\cdot)$ which resembles a multivariate Gaussian:

$$\begin{bmatrix} \mathsf{U}_1 \\ \mathsf{U}_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 3 \\ 7 \end{bmatrix}, \sigma^2 I \right)$$

The induced probability mass function π can be derived as follows:

$$\begin{cases} \pi(1 \succ 2) = \mathbb{P}_F(\mathsf{U}_1 > \mathsf{U}_2) \\ \pi(2 \succ 1) = \mathbb{P}_F(\mathsf{U}_2 > \mathsf{U}_1) \end{cases}$$